

Local All-Pass Geometric Deformations

Christopher Gilliam^{1b}, *Member, IEEE*, and Thierry Blu^{2b}, *Fellow, IEEE*

Abstract—This paper deals with the estimation of a deformation that describes the geometric transformation between two images. To solve this problem, we propose a novel framework that relies upon the brightness consistency hypothesis—a pixel’s intensity is maintained throughout the transformation. Instead of assuming small distortion and linearizing the problem (e.g. via Taylor Series expansion), we propose to interpret the brightness hypothesis as an all-pass filtering relation between the two images. The key advantages of this new interpretation are that no restrictions are placed on the amplitude of the deformation or on the spatial variations of the images. Moreover, by converting the all-pass filtering to a linear forward-backward filtering relation, our solution to the estimation problem equates to solving a linear system of equations, which leads to a highly efficient implementation. Using this framework, we develop a fast algorithm that relates one image to another, on a local level, using an all-pass filter and then extracts the deformation from the filter—hence the name “Local All-Pass” (LAP) algorithm. The effectiveness of this algorithm is demonstrated on a variety of synthetic and real deformations that are found in applications, such as image registration and motion estimation. In particular, when compared with a selection of image registration algorithms, the LAP obtains very accurate results for significantly reduced computation time and is very robust to noise corruption.

Index Terms—Geometric deformations, image registration, motion estimation, all-pass filters, spline and piecewise polynomial interpolation.

I. INTRODUCTION

THE estimation of the geometric transformation between two images is a problem that has many applications in image and video processing. For example, this problem is core to fundamental tasks such as image registration [3], image super-resolution [4] and video stabilization [5], and has been utilised in areas such as biology [6], medical imaging [7], remote sensing [8], fingerprint recognition [9]

Manuscript received December 22, 2016; revised June 11, 2017 and September 6, 2017; accepted October 5, 2017. Date of publication October 23, 2017; date of current version December 4, 2017. This work was supported in part by Huawei and in part by The Hong Kong Research Grants Council under Grant CUHK14200114. This paper was presented in part at the IEEE International Conference on Acoustics, Speech and Signal Processing, Brisbane, Queensland, Australia, and the IEEE International Conference on Image Processing, Quebec City, ON, Canada, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Weisheng Dong. (*Corresponding author: Christopher Gilliam.*)

C. Gilliam is with the School of Engineering, RMIT University, Melbourne, VIC 3000, Australia (e-mail: dr.christopher.gilliam@ieee.org).

T. Blu is with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: tblu@ee.cuhk.edu.hk).

This paper has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the author. The material includes a description of the synthetic data used in Section V and animations of the registration results. Contact dr.christopher.gilliam@ieee.org for further questions about this work.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2017.2765822

and fluid flow measurements [10]. Specifically, the problem comprises finding a transformation \mathcal{T} based on the variation of pixel intensities between the images. In this paper, rather than parametric registration (e.g. [11], [12]), we consider non-rigid image registration in which the geometric transformation can be characterised by a pixel-wise deformation field u . We solve this problem under the assumption that a pixel’s intensity remains constant under the deformation – known as the brightness consistency hypothesis [13], [14].

Mathematically, the brightness consistency hypothesis can be formulated as follows: given two 2D functions (i.e., images) $I_1(x)$ and $I_2(x)$, find a deformation field $u(x)$ that relates these two images through

$$I_2(x + u(x)) = I_1(x), \quad (1)$$

where $x = (x, y)^T$ is the pixel coordinates, $u(x) = (u_1(x), u_2(x))^T$ is a vector field and the transformation is $\mathcal{T}(x) = x + u(x)$. This formulation is, however, both restrictive and yet, at the same time, incomplete; the restrictiveness stems from the fact that (1) is usually not satisfied exactly (e.g., changes in illumination, sensor variation, atmospheric distortion etc.). Whereas, it is incomplete because, due to the dimensionality of $u(x)$ (equivalent to two images), there are many solutions to (1), most of them being meaningless; i.e., the problem is ill-posed [15]. Obtaining a meaningful deformation field is of particular importance in medical imaging where singularities and non-invertibility can be physically invalid [16].

In the registration literature, the approaches to solving (1) comprise two main aspects [7], [17]: the deformation model and the matching criteria. Starting with the first aspect, deformation models can be roughly split into two groups. The first are deformations derived from physical models such as elastic body models [18], fluid flow models [19] and diffusion based models (e.g. the Demons algorithm [20], [21]). These models are non-parametric in nature thus allowing a per pixel estimation for the deformation. The second group are deformation models derived from approximation theory. One popular example [22]–[26] is free form deformation whereby the deformation field is represented using basis functions, such as B-splines [27], at fixed integer grid positions. The advantage of these types of models is that they are capable of describing a wide range of transformations using a low/limited number of parameters [7]. Another approach is to assume the deformation can be approximated locally using a parametric model. A classic example is the Lucas and Kanade algorithm [13]: The brightness equation (1) is linearised using a first order Taylor expansion under the assumption that the deformation is small. Then, they approximated the deformation as constant

over a local region and estimated a constant vector u within such regions. Their method results in a per pixel estimate of the deformation field that can be non-rigid in nature. Note that this linearisation also formed the basis of early work on optical flow estimation [14].

Moving to the second aspect, the choice of matching criteria can also be divided into two groups. The first group comprises similarity measures based on the intensities of the images, e.g. the sum of squared difference or mutual information [28], [29]. More recently, Myronenko and Song [26] proposed minimising the number of basis functions that are required to code the error between the two images. In contrast, the second group focuses on matching features, such as SIFT points [30] or biologically relevant landmarks [7], between the images. In [8], this feature matching was used as an initialisation to a non-rigid registration based on mutual information. Along with these aspects, registration algorithms also place constraints on the invertibility of the deformation field [7]. Under the brightness hypothesis, an implicit assumption is that the deformation will warp image I_2 to image I_1 , which is true provided the transformation is invertible. This idea of an invertible transformation is either enforced directly using an inverse consistency constraint (e.g. local invertibility [31]) or by assuming a diffeomorphic transform [7], [32] in registration. For reviews of the state-of-the-art see [3], [7], [17].

The work presented in our paper is based on a novel formulation of the brightness consistency constraint, which changes the solving hypotheses quite substantially (e.g., arbitrary field amplitude, arbitrary spatial variations of the images). In contrast to the linearisation performed in [13], we express the left-hand side of (1) as a spatial filtering equation. The novelty does not lie in this structural property (exploited in the literature on image registration by [33]), but in the following observation: in the case where the field $u(x)$ is constant, this filtering formulation is exact and the spatial filter involved is *all-pass* [1]. When the deformation field is not constant but nonetheless slowly varying, the all-pass filtering relation between the images I_1 and I_2 is still “locally” valid up to a very good approximation. Given this local all-pass filter, we then apply a simple formula which provides an accurate estimate of the local displacement vector.

The key algorithmic idea in our paper is based on a forward-backward representation of all-pass filters: when I_2 is obtained by all-pass filtering of I_1 , it can equivalently be stated that adequate “forward” filtering of I_1 results in “backward” filtering of I_2 . There are no limitations on the nature of this forward filter, and the backward filter is simply the reversed version of the forward one. Thus, by using this representation, we can convert the non-linear all-pass constraint into a linear one when applied to (1). Accordingly, assuming the deformation is constant within a local region, we can solve the all-pass estimation problem very efficiently by solving a linear system of equations.

A. Relation to Prior Art

There are three main links between our work and prior art. First, similar to Lucas and Kanade, we assume the deformation field is locally constant and formulate an algorithm based on

solving a linear system of equations. However, it is important to stress that our framework does not require the additional assumption of small displacement; it can handle arbitrary sized displacements. Consequently, we do not require the Taylor approximation used in [13]. Moreover, in spite of using the same local constancy hypothesis, we obtain a consistent estimate of the deformation which contrasts with the usually inconsistent results obtained when applying the algorithm in [13]. Note that other local parametric models have been considered, e.g. local affine models [34], [35].

The second link is the use of spatial filtering in relation to the brightness consistency hypothesis. In general, the aim of this filtering in the literature has been to modify the properties of the hypothesis [36]. For example, Gaussian pre-filtering is used to reduce its sensitivity to noise [37] whereas higher order image derivatives are used to enforce illumination invariance [37]. More recently, the authors in [38] and [39] considered the optimal design of the pre-filters used in the Taylor linearisation proposed in [13]. In contrast, we depart from this idea of modifying the brightness hypothesis via filtering. Instead, we propose re-formulating the hypothesis itself as an all-pass filtering relationship and extract the deformation field from the corresponding filter. This relationship is then localised such that we obtain a local all-pass filter for every pixel and hence a dense, locally varying, deformation.

The final link is to a group of algorithms that estimate the deformation field based on the phase content of the images. A classic example is the estimation of a global translation and rotation using the finite Fourier transform [40]. Taking this idea further, Fleet and Jepson [41] proposed a spatio-temporal filtering approach to estimate the deformation. Their technique involved estimating the spatio-temporal Fourier phase of a sequence of images, via a bank of Gabor filters, and then using the variation in this phase to estimate the deformation. In this paper, however, we estimate spatial filters, per pixel, between two images and extract the deformation field from these filters (do not require a pre-defined filter bank).

On a last note, it is worth discussing the difference between image registration and optical flow estimation. The registration problem focuses on estimating a geometric deformation between two images. This deformation can then be used to either align the images (e.g. motion correction, image fusion), measure the geometric distortion of the imaging system [17] or analyse the change between the images [8]. For all of these tasks, a key element is that the deformation will warp image I_2 to image I_1 , which is true provided the field is invertible (locally). In contrast, the optical flow problem focuses on the estimation of a 2D motion field that describes the movement of 3D objects that have been projected onto a 2D image. Accordingly, a major challenge is the estimation of motion discontinuities. These discontinuities result in non-invertibility which invalidates the brightness hypothesis. One such example is occlusions; the optical flow is undefined in occluded regions as no mapping exists that warps I_2 to I_1 [42]. Thus, although some similarities exist, image registration and optical flow estimation are substantially different problems [43]. For a review of modern optical flow estimation we refer the reader to [36], [44].

B. Main Contributions and Outline

Using the idea of local all-pass filtering, we present a fast filter-based algorithm for estimating smoothly varying deformations, which we term the Local All-Pass (LAP) algorithm. We evaluate the performance of this algorithm when solving (1) in both ideal conditions (i.e. images that exactly satisfy the brightness constraint) and when the images are corrupted by noise. In comparison to a selection of image registration algorithms, our algorithm exhibits the following advantages: a large improvement in accuracy when estimating a deformation in which the brightness constraint is exactly satisfied; a robustness to violations of the brightness consistency caused by noise corruption or illumination change; and a significant reduction in computation time. We further demonstrate the versatility and robustness of the LAP when estimating deformations that occur in real images. In the context of the algorithm's speed, it is worth mentioning that the LAP is implemented using only basic Matlab commands, i.e., no optimised C or pre-compiled code, and this code is made available online.¹

This paper builds upon preliminary ideas presented in the conferences papers [1], [2] in the following ways: 1) We bring together the theoretical concepts presented in [2] to demonstrate how all-pass filters occur naturally from the assumption of small displacement. 2) We expand upon the iterative implementation of the LAP to enable it to cope with violations of the brightness consistency hypothesis (e.g. noise). 3) We provide a thorough analysis of the performance of the algorithm under varying image texture, violations of (1) due to noise/illumination change, and when varying the rate of change of the deformation field. We also demonstrate the algorithm on a number of real applications. Note that a 3D extension of the algorithm has also been presented in [45].

Finally, our paper is structured as follows: Section II details the main ideas of our all-pass filtering framework. Next, Section III describes the adaptation needed for this framework to work on a local level. Then, Section IV is devoted to additional processing techniques that are applied to the LAP to improve its robustness under non-ideal conditions. Our algorithm is then evaluated in Section V and Section VI shows examples on real images. We conclude in the last section.

II. ALL-PASS FILTER FRAMEWORK

Instead of assuming the deformation is small (and that the images have limited spatial variations), we shall assume the deformation field is slowly varying such that it is locally constant. The central concept is that a constant displacement caused by a rigid deformation is equivalent to filtering with an all-pass filter. Our framework thus consists of first estimating this all-pass filter and then extracting the displacement information from the filter. In the following section, we detail the mechanics of these procedures for the simplest case: a rigid deformation field with constant displacement. We extend to locally constant, slowly varying, deformation in Section III.

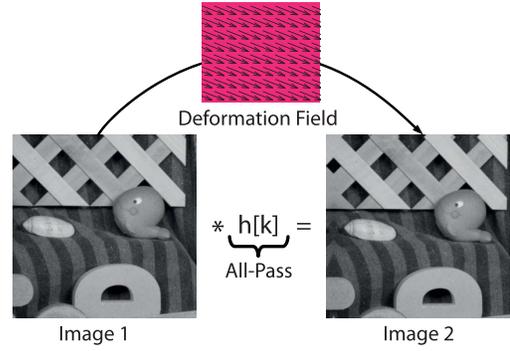


Fig. 1. Diagram illustrating the equivalence of a rigid deformation field with constant displacement and filtering using an all-pass filter $h[k]$. Note that $*$ is the convolution operator and $\mathbf{k} = [k, l]^T$ is the discrete pixel coordinates.

A. Main Principles of the All-Pass Framework

1) *Shifting is All-Pass Filtering*: We start by underlining the equivalence between a rigid deformation with constant displacement and all-pass filtering. Consider two arbitrary images (i.e., no hypothesis on their spatial variations), I_1 and I_2 , related by a constant displacement $\mathbf{u} = (u_1, u_2)^T$ as illustrated in Fig. 1. Assuming brightness consistency, we have that image I_2 is a shifted version of I_1 :

$$I_2(\mathbf{x} + \mathbf{u}) = I_1(\mathbf{x}). \quad (2)$$

In the frequency domain, this shifting relationship is equivalent to

$$\hat{I}_2(\boldsymbol{\omega}) = \hat{I}_1(\boldsymbol{\omega}) e^{-j\mathbf{u}^T \boldsymbol{\omega}}, \quad (3)$$

where \hat{I} represents the Fourier transform of the image I and $\boldsymbol{\omega} = (\omega_1, \omega_2)^T$ denotes the frequency coordinates. Now, if we define a filter h with a frequency response

$$\hat{h}(\boldsymbol{\omega}) = e^{-j\mathbf{u}^T \boldsymbol{\omega}}, \quad (4)$$

then we have that I_2 is a filtered version of I_1 and the filter in question has the following properties:

- **Separable**: $\hat{h}(\boldsymbol{\omega}) = \hat{h}_1(\omega_1) \hat{h}_2(\omega_2)$, where \hat{h}_1 and \hat{h}_2 are two 1D filters;
- **Real**: $\hat{h}(\boldsymbol{\omega}) = \hat{h}^*(-\boldsymbol{\omega})$, where \hat{h}^* represents the complex conjugate of \hat{h} , hence the impulse response is real-valued;
- **All-Pass**: $|\hat{h}(\boldsymbol{\omega})| = 1$.

An important observation from this analysis is that the information relating to the deformation field is contained within the phase of the all-pass filter h . Therefore, we propose estimating the filter h as a proxy for the complex exponential in (3).

To estimate this filter h , we need to move to the discrete domain. Thus, assuming ideal sampling with a sinc kernel, we obtain a digital version of the filter h which is, again, all-pass, and equal to the complex exponential (4) over the cyclic frequency range $\boldsymbol{\omega} \in]-\pi, \pi] \times]-\pi, \pi]$. Note that the other properties of this digital filter are the same as those found for the continuous version.

¹<https://sites.google.com/site/cwsgilliam/LAP>

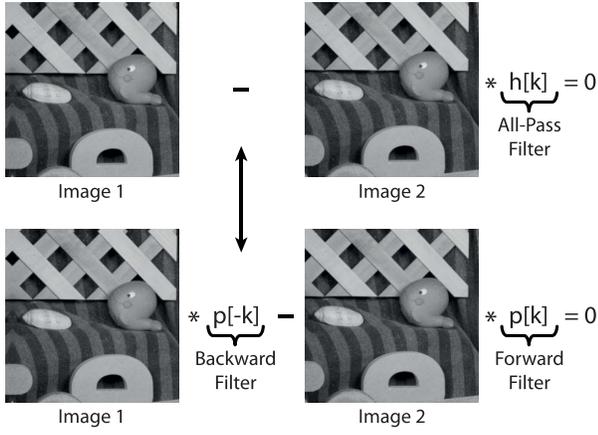


Fig. 2. Diagram illustrating the equivalence between two filtering operations in the pixel domain. This equivalence is due to the structure of the all-pass filter $h[k]$ defined in (5).

2) *Rational Representation of All-Pass Filters:* The $(2\pi, 2\pi)$ -periodic frequency response $\hat{h}(\omega)$ of any digital all-pass filter can always be expressed as the ratio of the DFT of two filters that have the same modulus, but opposite phase. More specifically, for a real all-pass filter

$$\hat{h}(\omega) = \frac{\hat{p}(e^{j\omega})}{\hat{p}(e^{-j\omega})}, \quad (5)$$

where $\hat{p}(e^{j\omega}) \stackrel{\text{def}}{=} \hat{p}(e^{j\omega_1}, e^{j\omega_2})$ is the forward and $\hat{p}(e^{-j\omega})$ the backward version of a real digital filter p . To verify this property, it suffices to choose $\hat{p}(e^{j\omega}) = \exp(j \arg(\hat{h}(\omega))/2)$ and to use the Hermitian symmetry that characterizes real filters; i.e., $\hat{h}^*(-\omega) = \hat{h}(\omega)$.

The importance of the representation in (5) is that the all-pass filtering relation between two images, I_1 and I_2 , can be transformed into a forward-backward filtering relation involving p . In other words, as illustrated in Fig. 2, the filtering operation performed by h can be expressed linearly as a function of p :

$$I_2[k] = h[k] * I_1[k] \Leftrightarrow p[-k] * I_2[k] - p[k] * I_1[k] = 0, \quad (6)$$

where $\mathbf{k} = [k, l]^T$ is the discrete pixel coordinate. Therefore, estimating the forward filter p in (6) is equivalent to estimating an all-pass filter that approximates the filter h in (4).

3) *Linear Approximation of the Forward Filter and its Estimation:* Using a standard signal processing technique, we approximate the filter p as a linear combination of a few fixed, known real filters p_n . In other words, we propose a filter basis representation:

$$p_{\text{app}}[k] = \sum_{n=0}^{N-1} c_n p_n[k], \quad (7)$$

where N denotes the number of filters in the basis and c_n are the coefficients of the filters. The benefit of this representation is that the estimation of the all-pass filter h is reduced to determining the N linear coefficients $\{c_n\}_{n=0, \dots, N-1}$. Note that, although we lose (in general) the separability property by making this approximation, we still preserve the real, all-pass properties of the filter h .

A consequence of the linear approximation of the filter p by (7) is that its estimation can be achieved using a straightforward mean square minimization of (6) over the coefficients c_n . Such a minimization is equivalent to solving an $N \times N$ linear system of equations, which, provided N is small, is very fast and efficient.

4) *Choosing a Good Filter Basis:* Having formulated the approximation of the all-pass filter, the question now becomes what type of filters should be used in (7)?

A naive choice is the canonical representation of finite impulse response (FIR) filters, supported in a square of side $2R+1$: $\hat{p}_n(e^{j\omega}) = e^{-j\omega_1 k_n} e^{-j\omega_2 l_n}$ where k_n, l_n are all the integers in $[-R, R]$. However limiting the support of the forward filter implies limiting the displacement of the deformation to a maximum of R pixels. Accordingly, the canonical FIR representation quickly becomes too expensive for estimating large scale deformations (e.g., for a deformation of displacement R pixels then $N \propto R^2$). Thus, we require a filter basis that has the two following properties: 1) Compactness— N is small compared to the number of pixels in the image so as to avoid overfitting and find a meaningful solution to the estimation problem. 2) Independence— N is independent of the size of displacement R so the filter basis can easily be scaled to estimate large displacement.

Our first approach to constructing such a filter basis was empirical; given an image I , we shifted the image using a set of known displacements vectors, whose amplitude remained the same but the direction varied from 0 to 360 degrees, and estimated the corresponding FIR filter p that satisfied (6). Once completed, we performed singular value decomposition (SVD) on the set of filters and took the first N eigenvectors, corresponding to the largest eigenvalues, to be the filter basis p_{app} . Thus, by choosing a small value for N we satisfy the compactness property. However, although this approach resulted in an optimum basis for I and that it seemed to be largely independent of the images [1], its disadvantage was that the basis needed to be recomputed if the size of the filters changed (i.e. it did not satisfy the independence property).

Interestingly however the filters obtained from this empirical method were generally well approximated using a Gaussian filter and its derivatives. Accordingly, we propose using a filter basis that spans the derivatives of a Gaussian function. To avoid having to estimate too many filter coefficients, we shall consider only two situations: i) Setting $K = 1$ (i.e. using only the first differentials), which equates to $N = 3$ filters. ii) Setting $K = 2$ (allowing both first and second differentials), which extends the basis to $N = 6$ filters. The discrete filters in question are

$$K = 1 \left\{ \begin{array}{l} p_0[k] = \exp\left(-\frac{k^2 + l^2}{2\sigma^2}\right) \\ p_1[k] = k p_0[k] \\ p_2[k] = l p_0[k] \\ p_3[k] = (k^2 + l^2 - 2\sigma^2) p_0[k] \\ p_4[k] = kl p_0[k] \\ p_5[k] = (k^2 - l^2) p_0[k] \end{array} \right\} K = 2 \quad (8)$$

where $\sigma = (R + 2)/4$ and R is the—integer—half-support of the filters. An important advantage of these filters is that they can easily be scaled to estimate larger deformations by changing the value of R . Thus the Gaussian filter basis satisfies both the compactness and independence properties that we require.

Now, although we arrived at this Gaussian filter basis empirically, there is a theoretical foundation to underpin this choice of basis. In [2], we examined the quality of approximation obtained when approximating the brightness consistency using (6). In particular, by using a Padé approximation, rather than a Taylor approximation, we showed that to obtain an approximation order of L (which is always even) the basis filters should comprise the derivatives, up to order $K = L/2$, of an isotropic function. As a consequence, by using $K = 1$ the filter basis can achieve an approximation order of 2 (see Section II-B below) -the equivalent approximation order using a Taylor expansion would require second order derivatives- and using $K = 2$ results in an approximation order of 4.

5) *From All-Pass Filters to Deformation Vectors*: The final piece of the framework is to retrieve the deformation from the estimated all-pass filter h_{est} . Since we expect the frequency response of the estimated filter, $\hat{h}_{\text{est}}(\omega)$, to be close to (4), we propose the following formula

$$u_{1,2} = j \left. \frac{\partial \log(\hat{h}_{\text{est}}(\omega))}{\partial \omega_{1,2}} \right|_{\omega_1=\omega_2=0}.$$

Given the frequency response of the all-pass filter in (5), the above formula has the following simple expression in terms of the impulse response of the filter p :

$$u_1 = 2 \frac{\sum_k k p[k]}{\sum_k p[k]} \quad \text{and} \quad u_2 = 2 \frac{\sum_k l p[k]}{\sum_k p[k]}, \quad (9)$$

where the summations are over all the multi-indices $\mathbf{k} = (k, l) \in \mathbb{Z}^2$. The above expressions were shown to be very accurate in [1].

B. Relation with Lucas-Kanade Approach - All-Pass Filters from Small Displacement Hypothesis

For this subsection only we shall assume the small displacement hypothesis and show how all-pass filters also naturally arise from this viewpoint. Using a first order Taylor expansion, the brightness consistency constraint can be replaced by an approximation where the deformation field coordinates are involved linearly [13]

$$I_2(x) = I_1(x - u) \approx I_1(x) - u^T \nabla I_1(x), \quad (10)$$

where $\nabla = [\partial/\partial x, \partial/\partial y]^T$. The advantage of the small displacement hypothesis is that the resulting approximation holds even if u is (possibly fast) varying. When u is constant, this equation is equivalent to $I_2 = h_0 * I_1$ where $h_0(x)$ is a continuous filter whose Fourier transform is given by $\hat{h}_0(\omega) = 1 - ju^T \omega$. Although this filter is not all-pass, it is the first order approximation (assuming u is small) of the all-pass filter (4).

Next, let us consider a symmetrical variant of the brightness consistency constraint -both images are shifted towards each

other by half the full displacement- and perform a first order expansion on both images:

$$\begin{aligned} I_2\left(x + \frac{u}{2}\right) &= I_1\left(x - \frac{u}{2}\right) \\ &\Downarrow \\ I_2(x) + \frac{1}{2}u^T \nabla I_2(x) &\approx I_1(x) - \frac{1}{2}u^T \nabla I_1(x) \end{aligned} \quad (11)$$

Again, this approximation holds even if u is not constant. Now, when u is constant, the relation between I_1 and I_2 is still a spatial convolution $I_2 = h_1 * I_1$ however the filter involved is all-pass

$$\hat{h}_1(\omega) = \frac{1 - ju^T \omega/2}{1 + ju^T \omega/2}.$$

This is, in fact, a second order Padé approximation² of the all-pass filter (4), as we have shown in [2]. Moving this equation to the discrete domain unveils the formal identity with (6).

Accordingly, by using the above interpretation, we observe that attempting to minimise (10) directly, as in [13], is equivalent to trying to find any filter that fits I_1 to I_2 . In contrast, performing the same minimisation on (11) is constrained so that only all-pass filters are considered. Consequently, the result is more likely to equate to a shifting operation (i.e. a valid deformation field).

This interpretation can be further developed so as to build approximations of the brightness consistency constraint that have increasing orders of accuracy. We have investigated this generalization in [2] by using a Padé approximation of the complex exponential function (4) with equal numerator and denominator order [46].

III. ESTIMATING SMOOTHLY VARYING DEFORMATIONS

We now relax the restriction of global constancy and allow the deformation field to vary in a slow, smooth manner. To estimate such a deformation, we propose a local adaptation to our framework based on the idea that the deformation can be considered as locally constant. Therefore, we can relate a local region in one image, I_1 , to the same corresponding region in another image, I_2 , using an all-pass filter; an illustration is shown in Fig. 3. Accordingly, we have a Local All-Pass (LAP) algorithm to estimate the deformation field. The mechanics of this algorithm are detailed as follows.

A. Obtaining Local All-Pass Filters

In order to estimate a local all-pass filter per pixel, we formulate a local version of the relation in (6) over a window \mathcal{W} of size $(2W + 1)$ by $(2W + 1)$ pixels. This new relation is termed the *local all-pass equation* and defined as follows:

$$p_{\text{app}}[k] * I_1[k] = p_{\text{app}}[-k] * I_2[k], \quad \text{where } k \in \mathcal{W}, \quad (12)$$

where $p_{\text{app}}[k]$ is the approximation of the FIR filter using the Gaussian filter basis. Note that the half-support of p_{app} is limited to $R \leq W$.

Now, to estimate the local filter $p_{\text{app}}[k]$, we opt to minimize, in the L_2 sense, the difference between the left and right sides

²The error between the lhs and the rhs of the approximation in (11) is $o(\|u\|^2)$.

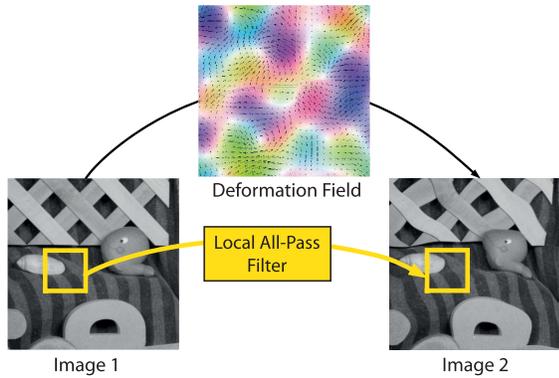


Fig. 3. Diagram illustrating that a smoothly varying deformation field between image 1 and image 2 can be approximated in a small region using a local all-pass filter. In this small region, the deformation between image 1 and 2 is equivalent to a convolution with an all-pass filter.

of (12) under the constraint that p satisfies (7); i.e., if we use the notations $\check{q}[k] \stackrel{\text{def}}{=} p[-k]$ and $\langle J \rangle_{\mathcal{W}} \stackrel{\text{def}}{=} \sum_{k \in \mathcal{W}} J[k]$, then we have to minimize the following expression, quadratic in c_n

$$\min_{\{c_n\}} \left\langle |p_{\text{app}} * I_1 - \check{p}_{\text{app}} * I_2|^2 \right\rangle_{\mathcal{W}} \quad (13)$$

where

$$p_{\text{app}}[k] = p_0[k] + \sum_{n=1}^{N-1} c_n p_n[k].$$

The local filter obtained from this minimization corresponds to the central pixel in \mathcal{W} . An interesting point raised in the review process of this paper is that the estimation of these coefficients could be posed using other penalty terms. For example, Senst *et al.* [47] proposed a modern update for the Lucas-Kanade algorithm that used robust penalty terms when estimating the displacement vector. A similar approach could be considered in (13) however we leave that as a topic for future research.

Once we have obtained an estimate of the filter from the minimization in (13), we can then shift the window \mathcal{W} and repeat the process for a new local all-pass filter. Based on this concept, we have the LAP algorithm - estimate a local all-pass filter for every pixel in the image, then, using these filters, extract an estimate of the deformation field according to (9). The process is briefly summarized in Algorithm 1. Notice that there are three parameters for the LAP algorithm: W , the half-width of the window \mathcal{W} , R , the half-support for the filters and K , the maximum derivative order of the filter basis.

Although similar to the formulation by Lucas and Kanade [13], this LAP algorithm has two important advantages: first, as established in [2], for small displacements the order of approximation is twice that of the Taylor expansion (10). Second, by changing the size of the window, the algorithm can easily scale to match the displacement of the deformation (remember that R acts as the upper bound on the displacement). Therefore, the LAP algorithm is not restricted to deformations of small displacements.

Algorithm 1 Local All-Pass (LAP) Estimation of a Deformation Field

Inputs: Images I_1 and I_2 , window size W , filter size R and number of filters N

- 1 Initialisation: Given a value of R and N , generate the basis filters described in (8). Note that $\sigma = (R + 2)/4$;
 - 2 Filter Estimation: Using the filter basis, solve the minimisation defined in (13) for each pixel in I_1 for a given local window of size $(2W + 1)$ by $(2W + 1)$, details in Section III-A1;
 - 3 Extraction: Using the set of filter coefficients, calculate the deformation for each pixel in the image using (9);
-

1) *Implementation Details:* The minimum of (13) satisfies the following equations, linear in c_n :

$$0 = \langle (p_n * I_1 - \check{p}_n * I_2)(p_0 * I_1 - \check{p}_0 * I_2) \rangle_{\mathcal{W}} + \sum_{n'=1}^{N-1} c_{n'} \langle (p_n * I_1 - \check{p}_n * I_2)(p_{n'} * I_1 - \check{p}_{n'} * I_2) \rangle_{\mathcal{W}},$$

for $n = 1, 2, \dots, N - 1$.

Now, the $\langle \cdot \rangle_{\mathcal{W}}$ operator is nothing else than a filter (the invariant shape of the region \mathcal{W} is shifted by integer values). Hence, the coefficients of $c_{n'}$ in the equations above can be computed by performing convolutions (fast). Finally, solving for the c_n at every pixel can be obtained, either using an explicit formula, or by Gauss elimination, both approaches involving only pointwise operations. These observations are the key to a fast implementation of the LAP algorithm.

B. Comparison with the Lucas-Kanade Algorithm

In contrast to Lucas-Kanade [13], the LAP algorithm described above is able to provide a very consistent and accurate deformation. To demonstrate this performance, we show in Fig. 4 the estimate of a constant deformation field obtained using the raw LAP, no pre/post-processing, and the equivalent obtained by the Lucas-Kanade algorithm [13]. To allow a fair comparison, the displacement of the deformation is 1 pixel, thus ideally suited for the Lucas-Kanade algorithm, and we set $W = R = 2$ for the LAP. In terms of quantitative values, if we average over 100 random realisations of the deformation, the LAP achieves a mean absolute error of 0.039 pixels when $K = 1$, and 0.021 pixels when $K = 2$; equivalent to a 3.9% and 2.1% error in the estimate, respectively. In contrast, the Lucas-Kanade algorithm obtains a mean absolute error of 0.403 pixels, which is equivalent to a 40.3% error in the estimate.

C. Poly-Filter Extension

Although the LAP algorithm is able to estimate large deformations directly, it requires a filter basis with large support to do so. This is equivalent to assuming large regions of the deformation are very slowly varying (i.e. constant), an assumption that is not likely to be always true. To overcome this issue, we use an iterative refinement in which the deformation field is estimated in a course-to-fine manner. However, unlike [13], [23], [48], we do not implement this refinement using

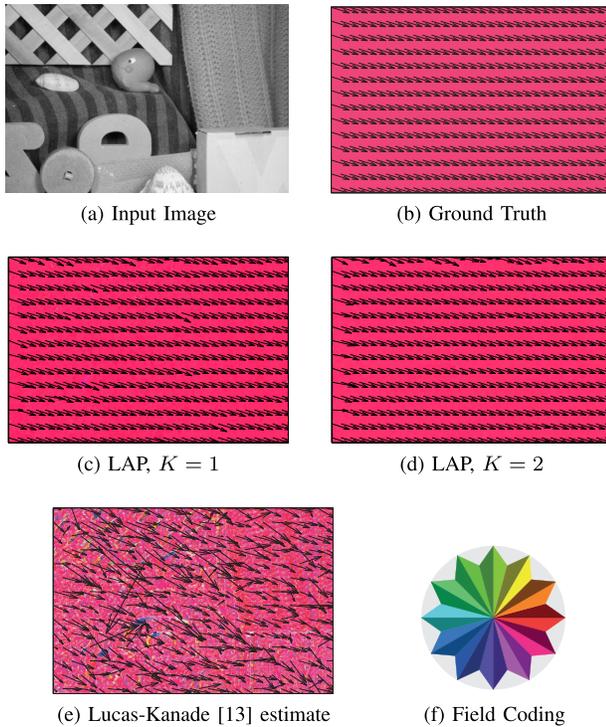


Fig. 4. Diagrams comparing the consistency of the raw LAP algorithm (without any pre/post processing) and the Lucas-Kanade algorithm [13] when estimating a constant, rigid, deformation field. The input image, (a), is synthetically deformed using the ground truth deformation shown in (b). The displacement of the deformation is 1 pixel. Note that (e) illustrates the colour coding for the deformation field (each colour represents a different direction of the deformation).

image pyramids, rather we only change the support of the filters using the parameter R , and continue to work on the full resolution image; large values of R allow the estimation of large slowly varying deformations whilst small values allow faster variation in the deformation. We term this extension the poly-filter LAP (PF-LAP). Note that, like all coarse-to-fine schemes, this approach will encounter difficulties if there are very localised large displacements in the deformation field.

The PF-LAP operates as follows: we first set the half-support of the filters, R , to its maximum value. For this filter size, we calculate the coarse, slowly varying, deformation field using the LAP algorithm. Next, we warp image I_2 closer to image I_1 using the estimate of the deformation. Then, we repeat the process with smaller filter sizes, each time adding the result to the previous estimate of the deformation, until we obtain a complete estimate that covers both slow and fast variations in the deformation field. To avoid enhancing errors in the deformation estimate (e.g., caused when the system of equations defined in (13) is ill-conditioned), we apply a post-processing procedure at each iteration. The post-processing comprises first detecting and replacing the errors using an inpainting procedure [49], [50], and then smoothing the resulting deformation field using a low-pass filter. The errors are identified in two ways: 1) if the displacement of the deformation estimate is greater than R , the size of the LAP filters; 2) if they are within W from the image boundary. The whole process is illustrated in Fig. 5 and

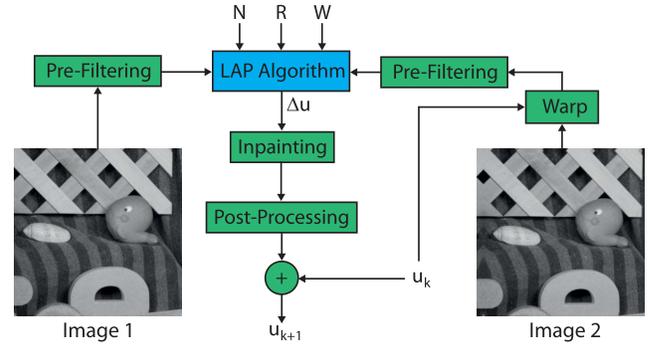


Fig. 5. Diagram illustrating the poly-filter extension to the LAP (termed PF-LAP). The LAP algorithm is summarised in Algorithm 1 and the processing blocks (Pre-Filtering, Warp, Inpainting and Post-Processing) are described in Section IV. Note that u_i is the deformation estimate at the i th iteration, Δu is the deformation increment, and N , R and W are respectively the number of filters, filter half-support and window half-support for the LAP.

Algorithm 2 Poly-Filter Extension to the LAP Algorithm

Inputs: Images I_1 and I_2 , number of filters N and vector of filter sizes r

- 1 Initialisation: Set $u_0 = \mathbf{0}$ and $I_2^{\text{shift}} = I_2$;
- 2 Noise Estimation: Estimate the noise in the images using method described in Section IV-B and set W_{limit} according to (14);
- 3 **for** $i = 1$ **to** Number of elements in r **do**
- 4 LAP Parameters: $R = r[i]$ and $W = \max\{R, W_{\text{limit}}\}$;
- 5 **for** $l = 1$ **to** Max. Iteration **do**
- 6 Pre-Filtering (OPTIONAL): Construct filter p_0 from (8) and obtain high-pass images I_1^{hp} and I_2^{hp} according to Section IV-A;
- 7 Estimation: Using N , R and W , estimate the deformation increment, Δu , between the images I_1^{hp} and I_2^{hp} using the LAP defined in Algorithm 1;
- 8 Post-Processing: Use inpainting procedure detailed in Section IV-D to remove errors in deformation increment, Δu . Then smooth Δu using Gaussian filter detailed in Section IV-E;
- 9 Update: Set $u_i = u_i + \Delta u$;
- 10 Warping: Warp I_2 closer to I_1 using u_i to obtain I_2^{shift} . If $R \leq 2$ pixels then warping is performed using cubic OMOMS interpolation, else it is performed using shifted-linear interpolation. See Section IV-C for details;
- 11 **if** $\text{PSNR}(I_1, I_{2,l}^{\text{shift}}) - \text{PSNR}(I_1, I_{2,l-1}^{\text{shift}}) > \epsilon$ **then**
- 12 | Terminate inner For loop;
- 13 **end**
- 14 **end**
- 15 **end**

summarised in Algorithm 2. The details of the image warping, inpainting and deformation smoothing are covered in the next section.

In terms of the actual implementation, we define a maximum value of the filter half-support, R_{max} , and then allow the values of R to descend in powers of 2 from R_{max} until $R = 1$ is reached. The exact value of R_{max} can either be determined based on a priori knowledge of the maximum displacement of the deformation (i.e. $R_{\text{max}} = \text{maximum displacement in pixels}$) or set such that the filter size does not exceed the size of the image being processed. Also, as the deformation estimate is smoothed at each iteration in the post-processing,

it can be beneficial to allow the algorithm to repeat iterations with the same size of filters. However, to avoid performing too many iterations, we limit the number, i.e. $Max. Iteration = 3$, and allow the algorithm to stop if the benefit is small. Note that the benefit is measured as the PSNR between the first image and the shifted version of the second image obtained using the deformation estimate.

Finally, we have to set the size of the local windows considered by the LAP algorithm (i.e., the parameter W). In noiseless conditions, this parameter should be set equal to the filter half-support R hence reducing in size as R decreases. If however the images are contaminated by noise then W should be limited in size, see Section IV-B for the precise limit. The reason for this is that at small window sizes the algorithm cannot distinguish between the deformation and the noise corruption. Thus, by increasing the window size, a larger linear system of equations, i.e., (13), can be built to combat the noise corruption. Note that the filter size is unchanged. A side effect of this approach to noise is that we lose some of the fine variations in the deformation field but as we shall demonstrate later this loss in accuracy is far less than current image registration algorithms.

IV. PRE- AND POST PROCESSING TOOLBOX

In this section, we present a toolbox of pre- and post-processing techniques that can be applied either on the images or on the estimate of deformation field in the PF-LAP.

A. Image Pre-Processing

If the brightness consistency hypothesis is satisfied then the PF-LAP can be used directly on the images. However, in practice, real images are likely to contain some change in illumination that invalidates the brightness consistency. To overcome this issue, several image pre-filtering techniques have been proposed ranging from simple Gaussian smoothing to non-linear structure-texture decomposition [51]. For the PF-LAP, we use two methods. The first is a simple histogram matching in which the histogram of I_2 is matched with I_1 . The second method is to perform a high-pass filtering operation on the images to reduce the effect of slowly varying illumination changes that are additive in nature. Specifically, we filter each image with the first filter (a Gaussian function) from the basis in (8), and then remove the smoothed image from the original: $I_i^{hp}[k] = I_i[k] - p_0[k] * I_i[k]$, where p_0 is the first basis filter and I_i^{hp} is high-passed version of the image I_i .

B. Image Noise Estimation

Along with varying illumination, the brightness consistency can also be invalidated if the input images are corrupted by noise. In this paper we shall assume this noise is additive white Gaussian in nature. Thus, to improve the LAP's robustness to such corruption, we limit the minimum window size W based on the amount of noise present in the image. Empirically, we obtain the following equation for the limit:

$$W_{limit} = \max(\lceil 38 - \text{PSNR}_{est}/2 \rceil, 1), \quad (14)$$

where PSNR_{est} is an estimate of the PSNR derived from the variance of the noise. Note that we use the median of the absolute deviation (MAD) of the highest frequency subband proposed in [52] to estimate the noise variance.

C. Image Warping

The iterative procedure of the PF-LAP algorithm requires applying a space-varying non-integer shift estimate to the image I_2 . To this end, it is necessary to build a continuous model of this image, which is achieved by interpolation. High-quality interpolation can be achieved using cubic-splines [27], [53], or Cubic OMOMS [54]. However, in order to ensure fast (speed similar to bilinear), but still high-quality interpolation (quality similar to Key's cubic interpolation [55]), we have opted for shifted linear interpolation [56] for all the LAP filters with a half-support $R > 2$ pixels. At finer filter resolution, i.e. $R \leq 2$ pixels, we use cubic-OMOMS interpolation.

D. Deformation Inpainting

The raw LAP algorithm is able to provide a very consistent deformation field. However, some values found are obviously erroneous because they exceed the maximal displacement expected, given the support of the forward filter. These values are then estimated based on their neighbours, using an isotropic diffusion that is very similar to the one proposed in [49], [50], and [57]. In essence, the valid (i.e., error free) neighbours of erroneous pixels are averaged to provide an estimate of these pixels. The process is then repeated (and refined for each erroneous pixel) until all erroneous pixels have been re-estimated. Note that motion inpainting has previously been proposed in [58] and [59]. Finally, simple replication of the neighbouring deformation values is used to fill in the boundary errors.

E. Deformation Smoothing

The aim of deformation smoothing is to remove any outlying estimates not previously identified in the inpainting procedure. Given our hypothesis of a slowly varying deformation field, we choose to smooth the estimate using a Gaussian filter. More specifically, given the local window half-size of W , we use a Gaussian filter that has $\sigma_s = 2W$ and a support of $4W$ by $4W$ pixels. Note that the filtering is performed assuming symmetric boundary conditions.

V. COMPARATIVE PERFORMANCE EVALUATION

In this section, we now evaluate the performance of the PF-LAP algorithm against the following selection of image registration algorithms: Matlab's Demons algorithm based on the implementation in [21]; intensity-based image registration using residual complexity minimisation from the Medical Image Registration Toolbox (MIRT) [26]; elastic registration using a cubic B-spline free form deformation model implemented in ImageJ (bUnwarpJ) [24]; and mutual information-based registration using a multiresolution cubic B-spline deformation model implemented in Elastix [25]. We also

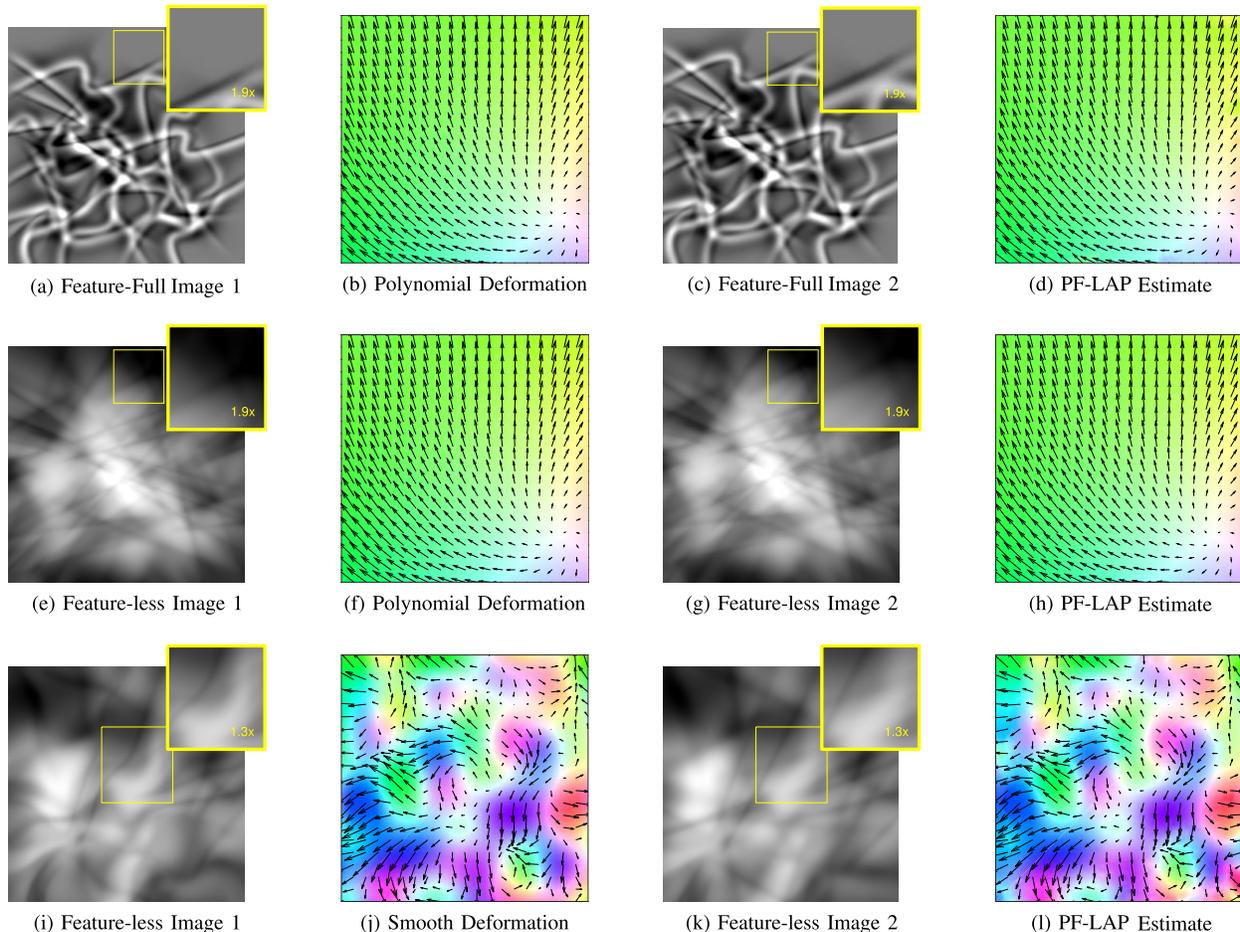


Fig. 6. Examples illustrating the synthetic ‘spaghetti’ images, the synthetic deformation fields and the corresponding estimate obtained by the PF-LAP. Parts (a) to (d) correspond to the feature-full images obtained when the thickness of the spaghetti is small whereas parts (e) to (h) correspond to the feature-less images obtained when the thickness of the spaghetti is large. Parts (i) to (l) correspond to the feature-less images under smoothly-varying deformation analysed in Section V-D. Note that the maximum displacement of the parametric deformation is 16 pixels and the smooth deformation is 25 pixels.

use, as a baseline, a modern implementation of Lucas and Kanade’s algorithm (LK) available in Piotr’s Computer Vision Matlab Toolbox [60]. The parameters for each algorithm are set according to their default values. For the LAP, we set $N = 3$ and choose R_{\max} to be the largest power of 2 that fits within the images (the minimum filter size is $R = 1$).

The performance of the algorithms is measured using three quantities: the median and mean absolute deformation error (in pixels) between the original deformation field and its estimate, E_{Med} and E_{Mean} , respectively; and the computation time of the algorithm. Note that, regarding the computation time, the algorithms are run on a desktop containing an Intel Core i7-5930K CPU @ 3.50 GHz with 64 Gb memory.

A. Generating Synthetic Data

To perform the evaluation, we generate synthetic images and deformation fields that exactly satisfy the brightness consistency in (1). The images are defined continuously as a summation of shifted, scaled and rotated cubic curves ($y = a_1x^3 + a_2x$), each with a certain length and thickness, which we term spaghetti. Examples of these spaghetti images are shown in Fig. 6 and their full description can be found

in the supplementary material. The advantage of this image description is twofold: first, by altering the parameters of the individual spaghetti, we can obtain varying types of image texture. For example, Fig. 6(a) and 6(e) demonstrate the variations obtained when changing just the thickness of the spaghetti. Secondly, by having a continuous description of I_2 , we can then automatically generate I_1 directly using (1); thus the images are independent of an interpolation method. Note that the size of each image is 301 by 301 pixels.

The synthetic deformation fields are defined continuously using a complex-valued function $f(z)$, where $z = x + jy$. Specifically, the field is obtained from the function as follows:

$$u(x) = [\text{Re}\{f(z)\}, \text{Im}\{f(z)\}]^T,$$

where $\text{Re}\{f(z)\}$ and $\text{Im}\{f(z)\}$ are respectively the real and imaginary parts of $f(z)$. The complex-valued function used in these synthetic simulations is a quadratic polynomial, $f(z) = b_1 + b_2 z + b_3 z^2$, which results in a smooth, slowly varying, deformation field. An example of a polynomial deformation field is shown in Fig. 6(b). Note that for these simulations we set the maximum displacement of the deformation field to be 16 pixels.

TABLE I

ERROR COMPARISON FOR THE PF-LAP AND A SELECTION OF IMAGE REGISTRATION ALGORITHMS UNDER IDEAL CONDITIONS

	Feature-full Images			Feature-less Images		
	E_{Med}	E_{Mean}	Time	E_{Med}	E_{Mean}	Time
PF-LAP	0.007	0.133	1.69	0.007	0.058	1.59
Demons [21]	0.053	0.180	2.81	0.026	0.121	2.69
bUnwarpJ [24]	0.038	0.152	5.95	0.053	0.106	6.82
Elastix [25]	0.096	1.162	956.7	0.068	0.159	958.4
MIRT [26]	0.086	0.956	30.30	0.023	0.239	29.60
LK [60]	0.105	0.525	0.02	0.323	1.245	0.02

* Results averaged over 10 random realisations of the deformation field (maximum displacement is 16 pixels)

** Bold values indicate the best results

B. Varying Image Texture

We start by analysing the performance of the algorithms in ideal conditions where our hypotheses - brightness consistency and smooth, slowly varying, deformation - are satisfied. Specifically, we evaluate the estimation of the quadratic deformation field, defined in the previous section, for two scenarios: 1) feature-full images in which the texture contains features such as edges, see Fig. 6(a); 2) feature-less images in which the texture varies slowly, see Fig. 6(e). These textures are obtained by generating identical spaghetti images, as discussed in Section V-A, and varying the thickness of the spaghetti; the feature-full images correspond to a thin spaghetti whereas the feature-less images correspond to a thick spaghetti, see the supplementary material for details. Note that each image contains 25 spaghetti objects and that the results are averaged over 10 random realisations of each type of image. Finally, as the brightness consistency is satisfied, no image pre-filtering is used in the PF-LAP.

The results of estimating the deformation fields with each algorithm are shown in Table I and examples of the estimates obtained by the PF-LAP are shown in Fig. 6. From the table, we observe that the PF-LAP is very accurate and consistently outperforms the other algorithms for both types of images. In particular, the PF-LAP is roughly 4 times more accurate than the next best algorithm in terms of median error for both types of images and roughly 2 times more accurate for the mean error on the feature-less images. Accordingly, our algorithm is able to estimate, almost exactly, a smoothly varying deformation field under the brightness consistency hypothesis. Also, unlike some of the other algorithms, its performance is independent of the image texture.

In terms of computation time, the PF-LAP took on average 1.64 seconds to compute the deformation field for the 301 by 301 pixel images; considerably faster than all of the algorithm with the exception of LK implementation [60]. Importantly, however, unlike [60] this computation time is achieved using only a Matlab implementation, without compiled mex-files. Also, due to the flexible nature of the PF-LAP, we have the option to further reduce the computation time by decreasing the parameter *Max. iteration* from 3 to 1 in Algorithm 2. In this case, the PF-LAP achieves a computation time of 0.6 seconds with only a small decrease in estimation accuracy;

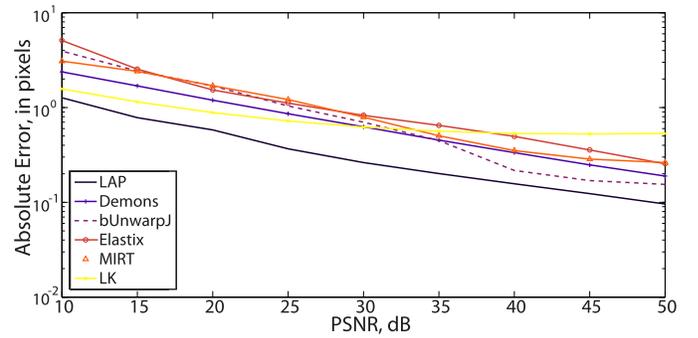


Fig. 7. Graph showing the mean absolute deformation error, E_{Mean} , for the PF-LAP and a selection of image registration algorithms obtained when the input images are corrupted by varying levels of additive white Gaussian noise.

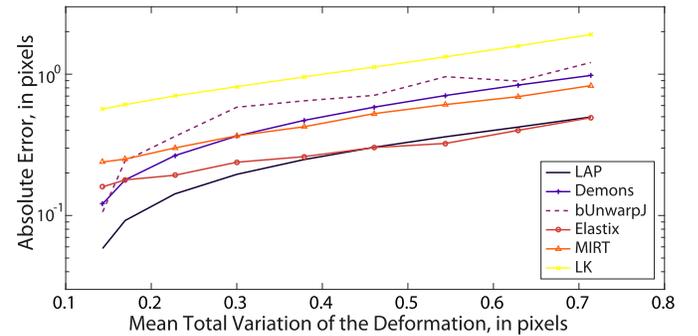


Fig. 8. Graph showing the mean absolute deformation error, E_{Mean} , for the PF-LAP and a selection of image registration algorithms obtained as the mean total variation of deformation increases. The equation of the mean total variation is given in (15) and it is measured in pixels.

$E_{Med} = 0.010$ pixels and $E_{Mean} = 0.150$ pixels for the feature-full images and $E_{Med} = 0.010$ pixels and $E_{Mean} = 0.150$ pixels for the feature-less images.

C. Violating the Brightness Consistency Hypothesis

In this section, we evaluate the robustness of the algorithms to violations of the brightness consistency hypothesis. Accordingly, we use the feature-less images, with quadratic deformations, from the previous section and corrupt the images with additive white Gaussian noise. The Peak Signal-to-Noise-Ratio (PSNR) varies from 50dB to 10dB and the results at each PSNR value are averaged over 10 random realisations.

The results of this evaluation are shown in Fig. 7. The figure shows the mean absolute deformation error, on a log scale, for the algorithms, see the supplementary material for the corresponding graph for the median error. The graph demonstrates that the PF-LAP is robust to violations of the brightness consistency caused by Gaussian noise. Specifically, our algorithm significantly outperforms the competing algorithms when the PSNR value decreases below 35dB; over 2 times more accurate than the next best algorithm. Along with these results, we have also observed that PF-LAP is robust to slowly varying deterministic violations of the brightness consistency (additive in nature). Using the image pre-filtering operation, we obtained an estimation accuracy of $E_{Med} = 0.036$ pixels and $E_{Mean} = 0.238$ pixels for the feature-less images where I_1 had undergone an illumination

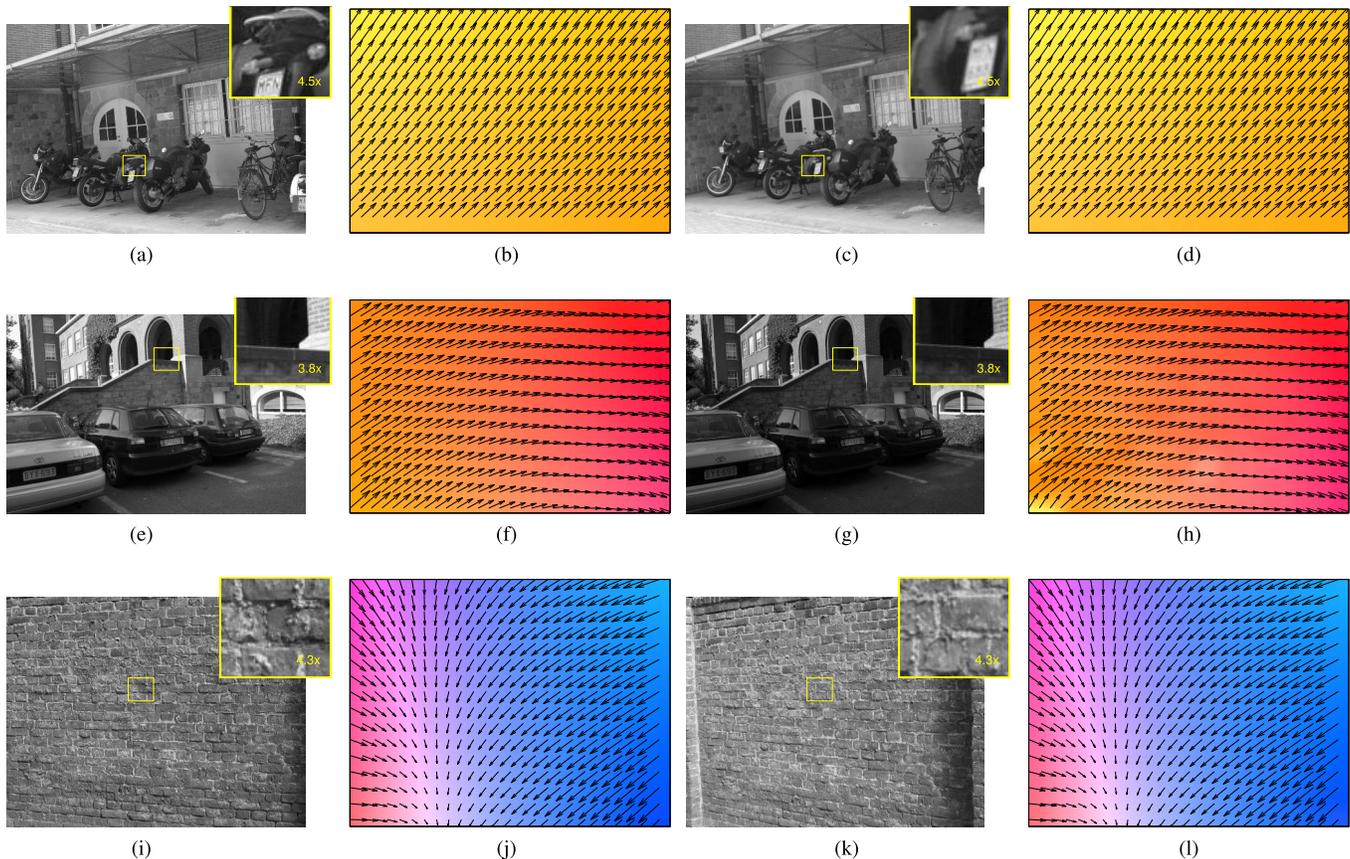


Fig. 9. Example of PF-LAP on three images from Oxford affine dataset [61]. The input images for Bikes are shown in parts (a) and (c), for Leuven in parts (e) and (g), and Wall in parts (i) and (k). The corresponding ground truth deformation of each image pair is shown in (b), (f) and (j), and the estimate obtained by the PF-LAP is shown in (d), (h) and (l). The maximum displacement for each deformation field is 46 pixels for Bikes, 7 pixels for Leuven and 119 pixels for Wall.

change characterised by a quadratic polynomial. Note that the next best algorithm achieved $E_{\text{Med}} = 0.778$ pixels and $E_{\text{Mean}} = 1.646$ pixels.

D. Probing the Local Constancy Assumption

We now evaluate the robustness of the algorithms to local variations in the deformation field (i.e. violations of the local constancy assumption). Specifically, we use the feature-less images from the previous tests but corrupt the quadratic deformations by adding a non-parametric, smoothly-varying, deformation field to them. An example of the resulting deformation is shown in Fig 6(j). The smoothly-varying deformation is obtained by low-pass filtering a randomly generated deformation field (each component of the field is drawn from a normal distribution with zero mean and a standard derivation of 1). The resulting deformation is then scaled so that it achieves a certain maximum displacement; increasing the maximum displacement equates to a faster varying deformation. For this evaluation the maximum displacement of the smoothly-varying deformation increases from 0 to 16 pixels. We measure how fast the deformation varies using the mean total variation, in pixels, of the deformation field:

$$\text{TV}_{\text{mean}} = \frac{1}{N_p} \sum_x \sqrt{\|\nabla u_1(x)\|_2^2 + \|\nabla u_2(x)\|_2^2}, \quad (15)$$

where N_p is the number of pixels in the images.

The results of this evaluation are shown in Fig. 8 and an example of the estimate obtained by the PF-LAP is illustrated in Fig. 6 (i) to (l). The graph in Fig. 8 shows the mean absolute deformation error, on a log scale, for the algorithms. Note that the first data point, on the left side, equates to the values in Table I for the Feature-less images (i.e. no smoothly-varying deformation field). The graph demonstrates that the PF-LAP is robust to violations of the local constancy assumption caused by a smoothly-varying deformation field. Also, with the exception of Elastix [25], the PF-LAP outperforms the comparison algorithms.

VI. APPLYING THE PF-LAP TO REAL IMAGES

In this section, we now demonstrate the versatility of our algorithm on real images and applications.

A. Image Alignment

We start by analysing the performance of the PF-LAP on a standard image processing task - image alignment. Specially, we use three pairs of images taken from the Oxford affine dataset [61]. This dataset provides pairs of images, which cover a range of situations e.g. blurring, varying illumination or change in viewpoint, and the corresponding homography that relates the two images. In more detail, we use the following images: 1) Bikes (size 1000 by 700 pixels), in which the second image has been corrupted via blurring;

TABLE II
ERROR COMPARISON FOR THE PF-LAP AND A SELECTION OF IMAGE REGISTRATION ALGORITHMS ON IMAGES FROM THE OXFORD AFFINE DATASET [61]

	Bikes			Leuven			Wall		
	E_{Med}	E_{Mean}	Time	E_{Med}	E_{Mean}	Time	E_{Med}	E_{Mean}	Time
PF-LAP	0.223	0.292	12.10	0.171	0.217	10.41	0.506	0.866	11.80
Demons [21]	0.458	0.702	15.09	0.243	0.395	11.83	1.383	21.66	13.16
bUnwarpJ [24]	0.220	0.308	8.54	0.189	0.229	11.59	10.57	30.28	30.69
MIRT [26]	0.726	4.228	129.9	0.363	0.813	87.76	0.571	1.874	110.4

* Bold values indicate the best results

2) Leuven (size 912 by 614 pixels), which in the second image has been corrupted by changing the illumination; 3) Wall (size 680 by 880 pixels), which contains a change in viewpoint and illumination. To assess the performance of the algorithm, we measure the accuracy of the estimated deformation field to that obtained from the ground truth homography provided with the images. All three pairs of images and the corresponding ground truth deformations are shown in Fig. 9. For comparison, we provide the results obtained when using Demons [21], bUnwarpJ [24] and MIRT [26]. Note that, for a fair comparison, histogram matching is applied to the Leuven and Wall images to combat the change in illumination before being processed by Demons, bUnwarpJ and MIRT.

The results of estimating the deformation fields with each algorithm are shown in Table II and the estimated deformations obtained by the PF-LAP are shown in Fig. 9. From the table, we observe that the performance of the PF-LAP is consistent across all three types of images. The deformation error increases slightly for the Wall images however the maximum displacement for this pair is 119 pixels whereas it is 46 pixels for Bikes and 7 pixels for Leuven. More generally, the PF-LAP outperforms the other non-rigid registration algorithms on the Leuven and Wall images, and obtains the lowest mean error on Bikes; it is 0.003 pixels off the lowest median error for Bikes achieved by bUnwarpJ. Similarly, in terms of computation time, the PF-LAP is faster for both the Leuven and Wall images. Further illustrations of these results are shown in the supplementary material.

B. Retinal Image Registration

We now analyse the performance on the PF-LAP when registering retinal images from the Fundus Image Registration (FIRE) dataset [62], available at <http://www.ics.forth.gr/cvrl/fire/>. The dataset comprises 134 retinal images pairs acquired with a Nidek AFC-210 fundus camera. The camera has a resolution of 2912 by 2912 pixels and a FOV of 45° both in the x and y dimensions. For this analysis we use the following two sets of images from the dataset: 1) Class S comprising 71 image pairs with an overlap greater than 75% and low anatomical differences between the images. 2) Class A comprising 14 image pairs with an overlap greater than 75% but with high anatomical differences. Note that we do not use the third group of images, Class P, as the overlap between the images is too small for non-rigid registration algorithms to obtain a meaningful result. An example of one of the Fundus image pairs is shown in Fig. 11.

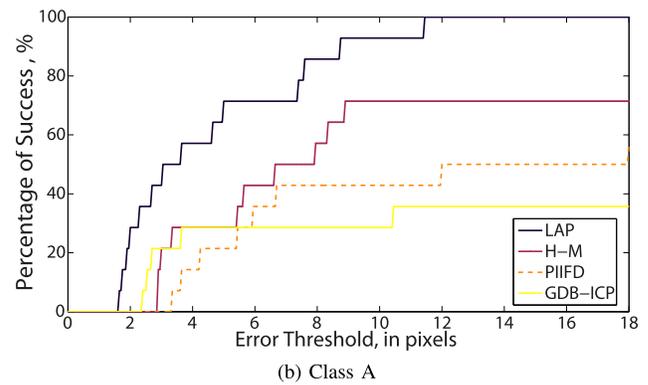
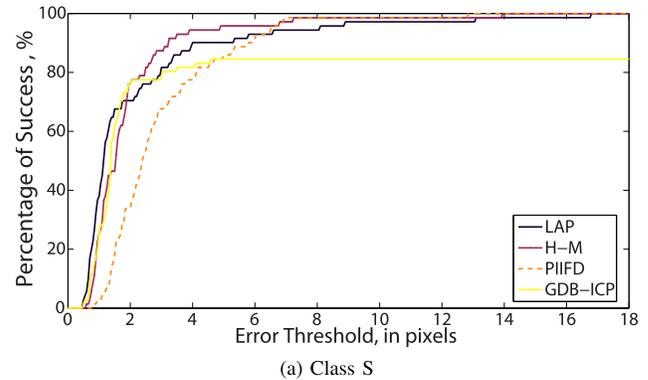


Fig. 10. Graphs showing the registration scores obtain by the LAP and other algorithms for the Class S and Class A retinal images from the FIRE dataset [62]. The scores indicate the percentage of images an algorithm successfully registered with an error less than a certain threshold. The error measures the distance between registered control points and their ground truth position. Note that there are 71 images in Class S and 14 in Class A.

To assess the performance of an algorithm, the dataset provides a sequence of control points in I_2 and the corresponding ground truth positions in I_1 . The accuracy of an algorithm is thus the displacement error, in pixels, between the registered control points and the ground truth versions. In [62], the displacement error per image pair is converted into a registration percentage score. The score is obtained by defining a registration as a success if the displacement error is less than a certain threshold and calculating the percentage of successful registrations for that threshold. An algorithm is then characterised by how the percentage score varies as error threshold increases. For comparison, we use the results provided by the dataset for the following algorithms: H-M [63], PIIFD [64] and GDB-ICP [65]. These algorithms

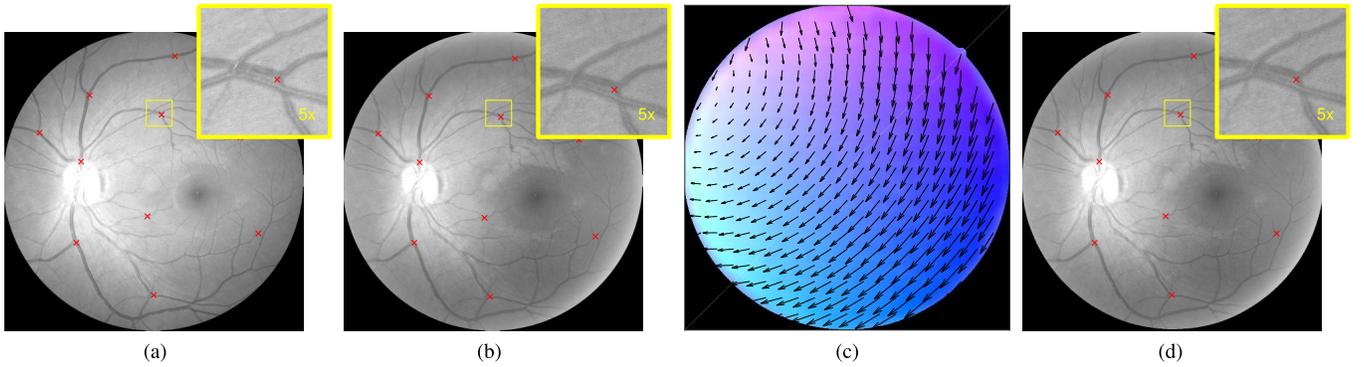


Fig. 11. Example of PF-LAP on a pair of retinal images from the Fundus Image Registration (FIRE) Dataset [62] (Class S). The input images are shown in parts (a) and (b), and the estimated deformation field from the LAP in (c). The registered version of image 2 obtained using this deformation is shown in (d). Note that the red crosses indicate the control points used to assess the performance of the registration. An animation illustrating the registration can be found in the supplementary material. Finally the images have been cropped to 2719 by 2719 pixels.

are all parametric in nature; they extract a set of feature points (Harris points in PIIFD, and SIFT in H-M and GDB-ICP) and then estimate a parametric model. A 2D parametric deformation model is used in PIIFD and GDB-ICP, whereas H-M estimates both the 3D camera pose and a 3D ellipsoidal model of the eye (12 parameters in total) to register the images. Thus these comparison algorithms are much more constrained and specialised than the PF-LAP—our aim is to demonstrate the versatile nature of the LAP rather than solving the specific FIRE problem. Note that the results for GDB-ICP are incomplete for both Class S and A, and the results for H-M and PIIFD are incomplete for Class A. Finally, due to the size of these images, we downsample them by 2 and then upscale the resulting deformation field.

The results of the registration are shown in Fig. 10, (a) corresponds to Class S and (b) to Class A, and examples of the registration obtained by the PF-LAP are shown in Fig. 11. The graphs in Fig. 10 show the percentage of successful registrations as the error threshold increases from 0 to 18 pixels. The graphs show that the PF-LAP outperforms the comparison algorithms for the Class A images and is very competitive for the Class S images. For example, on the Class S images, an error threshold of 1.5 pixels equates to a registration score of 68% for the PF-LAP compared to 46% for H-M [63], 62% for GDB-ICP [65] and 18% for PIIFD [64]. However, the added constraint of estimating just a parametric deformation enable the other algorithms to obtain a more accurate registrations on some of the images that exhibit greater illumination change and anatomical differences. This observation can be seen in the median and mean displacement errors: the PF-LAP obtains a median and mean error of 1.25 pixels and 2.17 pixels, respectively, on Class S and 3.32 pixels and 4.53 pixels on Class A. In comparison, the corresponding values for H-M [63] are 1.54 pixels and 1.95 pixels for Class S, and 7.28 pixels and 54.90 pixels for Class A. It would be interesting to see the results obtained if we combined a parametric model with the PF-LAP. Finally, in terms of computation time, on average the PF-LAP took 45.9 seconds per registration for Class S and 40.0 seconds per registration for Class A. The computation times for the other

algorithms were not provided. Further illustrations of these results are shown in the supplementary material.

C. Extension to 3D Motion Correction in MRI

Along with the previous examples, we have also presented, in [45], a 3D extension of the LAP and demonstrated using it to estimate and remove the respiratory motion present in three *in-vivo* MRI datasets. The estimated deformations obtained from the 3D LAP enable more accurate motion compensation of the MRI data than two competing non-rigid image registration algorithms: Elastix [25] and a 3D Demons algorithm [66]. In this case, the accuracy of the registration was measured by performing a lung segmentation on I_1 and comparing the volume of the lungs to the equivalent obtained from the registered version of I_2 . The comparison was measured using Dice coefficients [67]; a Dice coefficient of 1 indicates perfect overlap of the volumes whereas 0 indicates the opposite. Based on this measure, the 3D LAP achieved a mean Dice coefficient of 0.90 with a standard derivation of 0.01; in comparison Elastix achieved 0.87 with a standard derivation of 0.02 and Demons achieved 0.73 with a standard derivation of 0.05. Importantly for 3D imaging, the 3D LAP was also significantly faster than the other algorithms; it took 36 seconds to register a 256 by 256 by 72 volumetric image, which is roughly twice as fast as the next best algorithm.

VII. DISCUSSION AND CONCLUSIONS

We have presented a new approach to estimating a smoothly varying deformation field that does not require the small displacement hypothesis. The first ingredient of this approach is a reformulation of the brightness consistency hypothesis as local all-pass filtering; the local filter is used to relate a local region in one image to the same region in the other image. The deformation field is then extracted from the filter. The second ingredient is the linear transformation of the all-pass filtering using a forward-backward representation. Using this representation, we devised an efficient algorithm that required significantly less computation time than a selection of image registration algorithms. In synthetic experiments, we first

demonstrate that, without using any pre/post processing, our algorithm is able to obtain a tenfold increase in accuracy when compared to the standard Lucas and Kanade's algorithm [13], which also assumes the deformation is locally constant. Next, we show that the LAP outperforms a selection of image registration algorithms when tasked with estimate smoothly varying deformations that satisfy the brightness consistency. We then demonstrated that our algorithm maintains its competitive edge when faced with violations of the brightness consistency (e.g. image noise) and the local constancy assumption. Finally, we demonstrate the versatility of the LAP on real applications; our algorithm accurately, and quickly, aligned natural images taken from the Oxford dataset and registered retinal images from the Fundus Image Registration Dataset.

Accordingly, the proposed algorithm is well suited to retrieving smoothly varying flows, which is beneficial in non-computer vision applications such as registration in medical imaging, remote sensing and estimating fluid flows in meteorology. In particular, the key advantages of the LAP -its speed and accuracy- easily allow for the algorithm to be used as a basic tool within larger, more complex, tasks. For example we are currently in the process of incorporating a 3D version of the algorithm [45] into a MRI-based system that performs motion correction on PET images [68]. Another example is our recent work into iterative parametric registration [69]; the deformation obtained from the LAP is iteratively fitted to a parametric model. A final example is our use of the LAP algorithm to provide motion analysis in fluorescence microscopy images [70].

Finally, given the accuracy and speed of the LAP algorithm, it would be interesting to investigate how the LAP could be adapted to perform optical flow estimation. The key element here would be to address the challenge of estimating motion discontinuities, which, as mentioned in the introduction, cannot be modeled using the brightness consistency. One approach would be to follow the example of [71], [72] and use a two step process that involves first matching (i.e. running the original PF-LAP) and then using an edge-preserving interpolation method (e.g. EpicFlow [71]) to refine the matching. An alternative approach would be to investigate integrating the LAP equation (12) into the framework of cost-volume filtering [73], used in recent optical flow algorithms [74], [75].

ACKNOWLEDGMENT

We would like to thank Jinchun Yang, Yadong Lu, Roger Rui Chen (Huawei) and Pierre Moulin (UIUC) for useful discussions during the course of this project. We also thank the anonymous reviewers for their comments, which have improved the quality of this paper.

REFERENCES

- [1] C. Gilliam and T. Blu, "Local all-pass filters for optical flow estimation," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2015, pp. 1533–1537.
- [2] T. Blu, P. Moulin, and C. Gilliam, "Approximation order of the LAP optical flow algorithm," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 48–52.
- [3] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, 2003.
- [4] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Process. Mag.*, vol. 20, no. 3, pp. 21–36, May 2003.
- [5] C.-K. Liang, L.-W. Chang, and H. H. Chen, "Analysis and compensation of rolling shutter effect," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1323–1330, Aug. 2008.
- [6] L. Qu, F. Long, and H. Peng, "3D registration of biological images and models: Registration of microscopic images and its uses in segmentation and annotation," *IEEE Signal Process. Mag.*, vol. 32, no. 1, pp. 70–77, Jan. 2015.
- [7] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013.
- [8] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4328–4338, Jul. 2014.
- [9] L. Liu, T. Jiang, J. Yang, and C. Zhu, "Fingerprint registration by maximization of mutual information," *IEEE Trans. Image Process.*, vol. 15, no. 5, pp. 1100–1110, May 2006.
- [10] D. Heitz, E. Mémin, and C. Schnörr, "Variational fluid flow measurements from image sequences: Synopsis and perspectives," *Experim. Fluids*, vol. 48, no. 3, pp. 369–393, 2010.
- [11] P. Thevenaz, U. E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 27–41, Jan. 1998.
- [12] J.-M. Odobez and P. Bouthemy, "Robust multiresolution estimation of parametric motion models," *J. Vis. Commun. Image Represent.*, vol. 6, no. 4, pp. 348–365, Dec. 1995.
- [13] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. 7th Int. Joint Conf. Artif. Intell.*, vol. 2, 1981, pp. 674–679.
- [14] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, Aug. 1981.
- [15] T. Poggio, V. Torre, and C. Koch, "Computational vision and regularization theory," *Image Understand.*, vol. 317, p. 314, Sep. 1985.
- [16] W. R. Crum, T. Hartkens, and D. L. G. Hill, "Non-rigid image registration: Theory and practice," *Brit. J. Radiol.*, vol. 77, no. 2, pp. S140–S153, 2004.
- [17] M. Holden, "A review of geometric transformations for nonrigid body registration," *IEEE Trans. Med. Imag.*, vol. 27, no. 1, pp. 111–128, Jan. 2008.
- [18] C. Broit, "Optimal registration of deformed images," Ph.D. dissertation, Dept. Comput. Inf. Sci., Univ. Pennsylvania, Philadelphia, PA, USA, 1981.
- [19] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Image Process.*, vol. 5, no. 10, pp. 1435–1447, Oct. 1996.
- [20] J.-P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.
- [21] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, 2009.
- [22] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imag.*, vol. 18, no. 8, pp. 712–721, Aug. 1999.
- [23] J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1427–1442, Nov. 2003.
- [24] I. Arganda-Carreras, C. Sorzano, R. Marabini, J. Carazo, C. Ortiz-de Solorzano, and J. Kybic, "Consistent and elastic registration of histological sections using vector-spline regularization," in *Proc. Int. Workshop Comput. Vis. Approaches Med. Image Anal.* 2006, pp. 85–95.
- [25] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 196–205, Jan. 2010.
- [26] A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Trans. Med. Imag.*, vol. 29, no. 11, pp. 1882–1891, Nov. 2010.
- [27] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Process. Mag.*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

- [28] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.
- [29] P. Viola and W. M. Wells, III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, 1997.
- [30] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. L. Yuille, and Z. Tu, "Robust L_2E estimation of transformation for non-rigid registration," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1115–1129, Mar. 2015.
- [31] S. Y. Chun and J. A. Fessler, "A simple regularizer for b-spline nonrigid image registration that encourages local invertibility," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 1, pp. 159–169, Feb. 2009.
- [32] G. E. Christensen and H. J. Johnson, "Invertibility and transitivity analysis for nonrigid image registration," *J. Electron. Imag.*, vol. 12, no. 1, pp. 106–117, 2003.
- [33] H. S. Stone, M. T. Orchard, E.-C. Chang, and S. A. Martucci, "A fast direct Fourier-based algorithm for subpixel registration of images," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2235–2243, Oct. 2001.
- [34] D. Robinson and P. Milanfar, "Fast local and global projection-based methods for affine motion estimation," *J. Math. Imag. Vis.*, vol. 18, no. 1, pp. 35–54, 2003.
- [35] S. X. Ju, M. Black, and A. Jepson, "Skin and bones: Multi-layer, locally affine, optical flow and regularization with transparency," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Francisco, CA, USA, Jun. 1996, pp. 307–314.
- [36] D. F. P. Bouthemy and C. Kervrann, "Optical flow modeling and computation: A survey," *Comput. Vis. Image Understand.*, vol. 134, pp. 1–21, May 2015.
- [37] T. Brox, A. Bruhn, N. Papenberger, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Prague, Czech Republic, 2004, pp. 25–36.
- [38] M. Elad, P. Teo, and Y. Hel-Or, "Optimal filters for gradient-based motion estimation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 1, Dec. 1999, pp. 559–565.
- [39] D. Robinson and P. Milanfar, "Bias minimizing filter design for gradient-based image registration," *Signal Process., Image Commun.*, vol. 20, no. 6, pp. 554–568, 2005.
- [40] E. De Castro and C. Morandi, "Registration of translated and rotated images using finite Fourier transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 5, pp. 700–703, Sep. 1987.
- [41] D. Fleet and A. Jepson, "Computation of component image velocity from local phase information," *Int. J. Comput. Vis.*, vol. 5, no. 1, pp. 77–104, 1990.
- [42] A. Ayvaci, M. Raptis, and S. Soatto, "Sparse occlusion detection with optical flow," *Int. J. Comput. Vis.*, vol. 97, no. 3, pp. 322–338, 2012.
- [43] C. Xing and P. Qiu, "Intensity-based image registration by nonparametric local smoothing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 2081–2092, Oct. 2011.
- [44] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," *Int. J. Comput. Vis.*, vol. 92, no. 1, pp. 1–31, 2011.
- [45] C. Gilliam, T. Küstner, and T. Blu, "3D motion flow estimation using local all-pass filters," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Prague, Czech Republic, Apr. 2016, pp. 282–285.
- [46] G. Baker, Jr., and J. L. Gammel, "The padé approximant," *J. Math. Anal. Appl.*, vol. 2, no. 1, pp. 21–30, 1961.
- [47] T. Senst, V. Eiselein, and T. Sikora, "Robust local optical flow for feature tracking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 9, pp. 1377–1387, Sep. 2012.
- [48] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proc. Eur. Conf. Comput. Vis.*, Santa Margherita Ligure, Italy, 1992, pp. 237–252.
- [49] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proc. Int. Conf. Comput. Graph. Interaction Tech. (SIGGRAPH)*, 2000, pp. 417–424.
- [50] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2634–2645, Oct. 2010.
- [51] A. Wedel, T. Pock, C. Zach, H. Bischof, and D. Cremers, "An improved algorithm for TV- L^1 optical flow," in *Statistics Geometrical Approaches Vision Motion Analytical* (Lecture Notes in Computer Science), D. Cremers, B. Rosenhahn, A. Yuille, and F. Schmidt, Eds. Berlin, Germany: Springer, 2009, pp. 23–45.
- [52] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inf. Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
- [53] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 739–758, Jul. 2000.
- [54] T. Blu, P. Thévenaz, and M. Unser, "MOMS: Maximal-order interpolation of minimal support," *IEEE Trans. Image Process.*, vol. 10, no. 7, pp. 1069–1080, Jul. 2001.
- [55] R. Keys, "Cubic convolution interpolation for digital image processing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Atlanta, GA, USA, Dec. 1981, pp. 1153–1160.
- [56] T. Blu, P. Thévenaz, and M. Unser, "Linear interpolation revitalized," *IEEE Trans. Image Process.*, vol. 13, no. 5, pp. 710–719, May 2004.
- [57] M. Oliveira, B. Bowen, R. McKenna, and Y. S. Chang, "Fast digital image inpainting," in *Proc. Int. Conf. Vis. Imag. Image Process. (VIIP)*, Marbella, Spain, 2001, pp. 261–266.
- [58] Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H.-Y. Shum, "Full-frame video stabilization with motion inpainting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1150–1163, Jul. 2006.
- [59] C. Kondermann, D. Kondermann, and C. Garbe, "Postprocessing of optical flows via surface measures and motion inpainting," in *Pattern Recognition*. Berlin, Germany: Springer, 2008, pp. 355–364.
- [60] P. Dollár, *Piotr's Computer Vision MATLAB Toolbox, V3.50*. Accessed: May 20, 2017. [Online]. Available: <https://github.com/pdollar/toolbox>
- [61] K. Mikolajczyk et al., "A comparison of affine region detectors," *Int. J. Comput. Vis.*, vol. 65, no. 1, pp. 43–72, 2005.
- [62] C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, S. Douma, and A. A. Argyros, "FIRE: Fundus image registration dataset," *J. Model. Ophthalmol.*, vol. 1, no. 4, pp. 16–28, 2017.
- [63] C. Hernandez-Matas, X. Zabulis, and A. A. Argyros, "Retinal image registration through simultaneous camera pose and eye shape estimation," in *Proc. IEEE Conf. Eng. Med. Biol. Soc. (EMBC)*, Orlando, FL, USA, Aug. 2016, pp. 3247–3251.
- [64] J. Chen, J. Tian, N. Lee, J. Zheng, R. T. Smith, and A. F. Laine, "A partial intensity invariant feature descriptor for multimodal retinal image registration," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1707–1718, Jul. 2010.
- [65] G. H. Yang, C. V. Stewart, M. Sofka, and C.-L. Tsai, "Registration of challenging image pairs: Initialization, estimation, and decision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 11, pp. 1973–1989, Nov. 2007.
- [66] D.-J. Kroon and C. H. Slump, "MRI modality transformation in demon registration," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Boston, MA, USA, Jul. 2009, pp. 963–966.
- [67] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [68] T. Küstner et al., "MR-based respiratory and cardiac motion correction for PET imaging," *J. Med. Image Anal.*, vol. 42, pp. 129–144, Dec. 2017.
- [69] X. Zhang, C. Gilliam, and T. Blu, "Iterative fitting after elastic registration: An efficient strategy for accurate estimation of parametric deformations," in *Proc. IEEE Conf. Image Process. (ICIP)*, Sep. 2017, pp. 1492–1496.
- [70] J. Li, C. Gilliam, and T. Blu, "A multi-frame optical flow spot tracker," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Quebec City, QC, Canada, Sep. 2015, pp. 3670–3674.
- [71] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, "EpicFlow: Edge-preserving interpolation of correspondences for optical flow," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1164–1172.
- [72] Y. Hu, R. Song, and Y. Li, "Efficient coarse-to-fine patchmatch for large displacement optical flow," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 5704–5712.
- [73] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 3017–3024.
- [74] J. Lu, H. Yang, D. Min, and M. N. Do, "Patch match filter: Efficient edge-aware filtering meets randomized search for fast correspondence field estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Portland, OR, USA, Jun. 2013, pp. 1854–1861.
- [75] Y. Hu, R. Song, Y. Li, P. Rao, and Y. Wang, "Highly accurate optical flow estimation on superpixel tree," *Image Vis. Comput.*, vol. 52, pp. 167–177, Aug. 2016.



Christopher Gilliam received the degree (Hons.) from the Imperial College London, U.K., the M Eng. degree in electrical and electronic engineering in 2008, and the Ph.D. degree in electrical and electronic engineering in 2013. He is currently a Research Fellow with the School of Engineering, RMIT University, Australia. He was a Postdoctoral Fellow with the Image and Video Processing Group, Department of Electronic Engineering, The Chinese University of Hong Kong. His research interests include sampling theory and its application in image

processing in particular image-based rendering, sparse sampling involving signals with finite rate of innovation, motion estimation and registration for image sequences, in particular biological and medical imaging, and sensor management.



Thierry Blu (M'96–M'06–F'12) was born in Orléans, France, in 1964. He received the Dipl.-Ing degree from the Ecole Polytechnique, France, and Télécom Paris (ENST), Paris, in 1986 and 1988. In 1996, the Ph.D. degree in electrical engineering from ENST in iterated rational filterbanks, and applied to wideband audio coding.

From 1998 and 2007, he was with the Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne, Switzerland. He is currently a Professor with the Department of Electronic Engineering,

The Chinese University of Hong Kong. His research interests include wavelets, approximation and sampling theory, sparse representations, image denoising, biomedical imaging, optics, and wave propagation.

Dr. Blu was the recipient of two best paper awards from the IEEE Signal Processing Society in 2003 and 2006. He is also the co-author of a paper that received a Young Author Best Paper Award in 2009 from the IEEE Signal Processing Society. He has been a member of the IEEE Signal Processing Theory and Methods Technical Committee from 2008 to 2013, and an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 2002 to 2006, THE IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2006 to 2010, and the Elsevier Signal Processing from 2008 to 2011. He has been on the board of *Eurasip J. on Image and Video Processing*, since 2010, and a member of the IEEE Bio Imaging and Signal Processing Technical Committee.